# S1 Supplementary data

## S1.1 Incidence data

Monthly incidence data were available from the Massachusetts (MA) Department of Public Health (MDPH). The data were stratified by age (1-y breakdown for individuals <20 yo, 5-y breakdown for individuals ≥20 yo). For simplicity, we aggregated the data into 7 epidemiologically relevant age groups: infants 0–1 yo (i.e., [0, 1) yo); preschool children 1–5 yo ([1, 5) yo); school-aged children 5–10 yo ([5, 10) yo), 10–15 yo ([10, 15) y), and 15–20 yo ([15, 20) yo); adults 20–40 yo ([20, 40) y) and ≥40 yo. According to the MDPH[1], the case definition for pertussis in non-outbreak settings is: laboratory confirmation by culture in a patient with any cough illness; a cough illness lasting ≥2 weeks, with lab confirmation by serology in a person not vaccinated with a pertussis-containing vaccine in the 3 previous years; a cough illness lasting ≥2 weeks with 1 or more of the following: paroxysms of coughing, inspiratory whoop, or post-tussive vomiting, without other apparent cause, in an individual who has a positive PCR test; a cough illness lasting ≥ 2 weeks with one or more of the following: paroxysms of coughing, inspiratory whoop, or post-tussive vomiting, without other apparent cause, without appropriately-timed negative laboratory test, in an individual who is epidemiologically-linked to a lab-confirmed case.

A timeline of pertussis surveillance in MA is presented in table S1. Notably, serological testing became available in 1987 for individuals ≥11 y, leading to an immediate and substantial increase in the number of reported cases in 11–19 yo and ≥20 y (1). By contrast, introduction of PCR testing in January 2005 did not lead to noticeable increases in the number of reports in adolescents and adults (Fig. S1). In 2006, a booster reduced-dose of acellular pertussis vaccine combined with tetanus and diphtheria toxoids (Tdap) was recommended among adolescents 11 to 18yo (2). We therefore restricted our analysis to data during 1990–2005 (2005 included, $n_y = 16$ y of monthly data), a period of stable surveillance before the introduction of Tdap.

[1] http://tinyurl.com/p9mmhwv

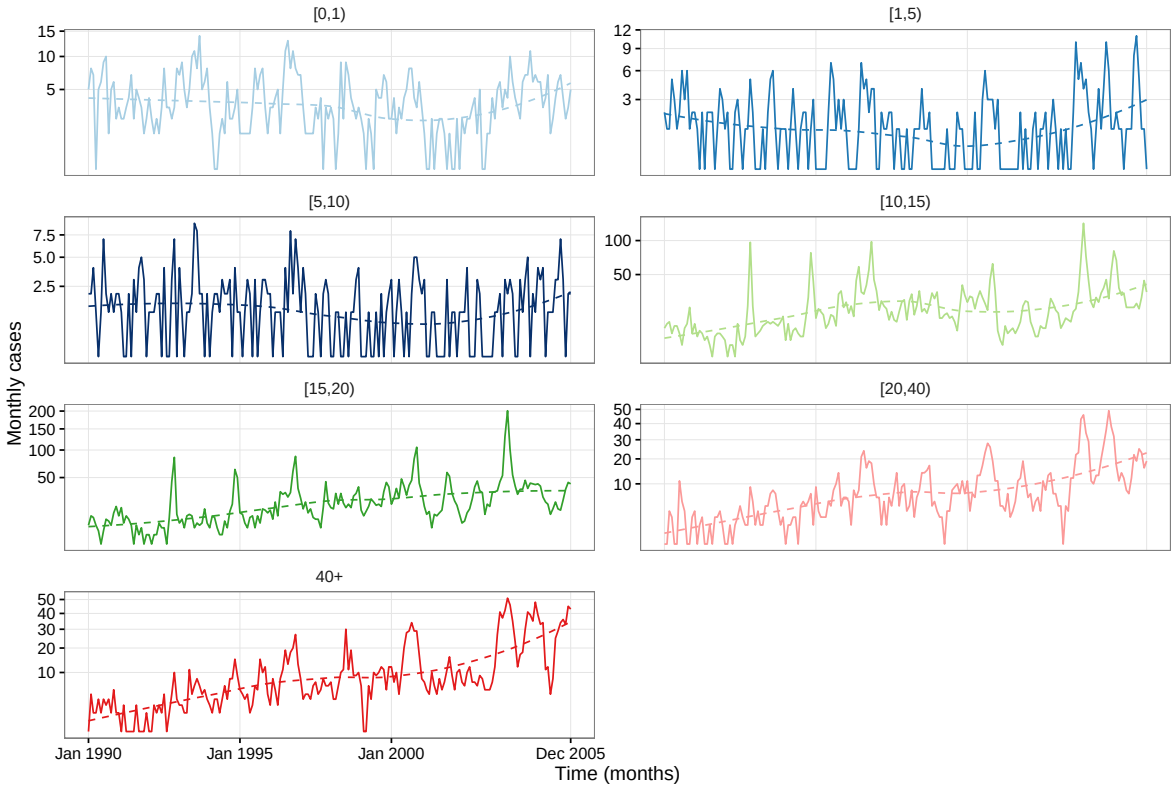


Figure S1: Monthly reported cases by age group during 1990–2005. Dashed lines represent smoothed data from loess regression (degree 2, span 0.75). For visual clarity, the $y$-axis is square-root transformed and differs between panels.

| Date | Event | Source |
|---|---|---|
| 1950 | Start of mass production of whole-cell pertussis (wP) vaccines | (1) |
| 1986 | Start of increased investigation of institutional outbreaks | (1) |
| 1987 | Start of serologic ELISA for ≥11 yo | (1) |
| 1989 | Explicit criteria for clinical diagnosis adopted | (1) |
| October 1996 | Switch to acellular pertussis vaccine (aP) Tripedia | (3) |
| January 2005 | Start of PCR testing | http://tinyurl.com/p9mmhwv |
| 2006 | Tdap booster dose recommended in 11–18 yo | (2) |

Table S1: Timeline of pertussis surveillance effort and of pertussis vaccination in MA.

## S1.2 Immunization data

Immunization levels of children entering kindergarten were available from from the MDPH, from school year 1975/1976 for children having received ≥4 doses and from school year 1995/1996 for children having received 5 doses (4). As shown in Fig. S2, the vaccine coverage was approximately constant during this period. Therefore, we assumed constant vaccine coverages $v_1 = 0.97$ and $v_2 = 0.93$, where $v_1$ represents the

vaccine coverage for the primary course and $v_2$ the conditional probability of having received a fifth dose given that 4 doses have been received.

In the absence of vaccination data before 1970, we made pragmatic assumptions based on available evidence. Although mass production of the wP vaccine began in 1950 in MA (1), it was already distributed across the US from 1940 (5). This is consistent with historical incidence data, which show that pertussis incidence in MA began decreasing in the 1940s and steeply declined after 1950 (6). We therefore assumed that vaccination had started in 1940 and that the vaccine coverage ramped up from 0 in 1940 to $v_1$ in 1955. We also assumed that the preschool booster dose began being administered in 1967 (7).

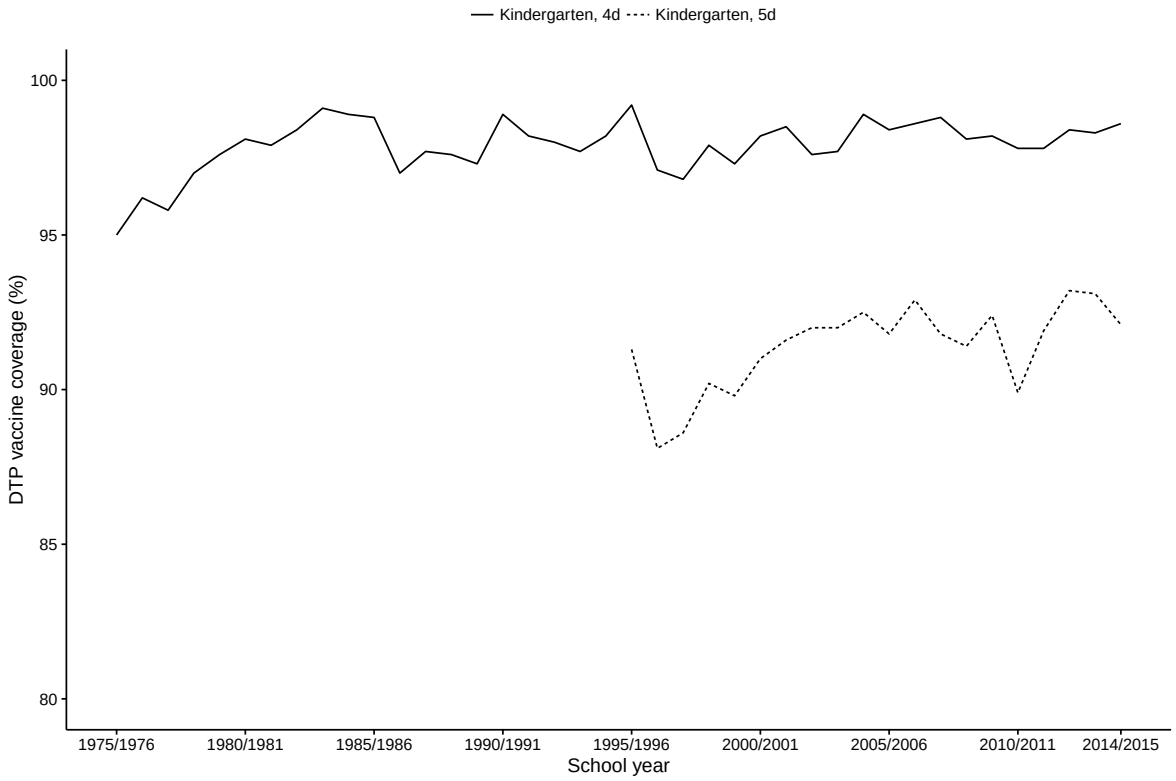


Figure S2: Pertussis vaccine coverage in MA. Source: MDPH immunization surveys (4).

## S1.3 Age-specific seasonality

Monthly incidence data by age group are presented in Fig. S1. To examine the age-specific seasonality, we computed, for each year $y = 1, \ldots, n_y$, the proportion of cases during month $m = 1, \ldots, 12$, $p_{m,y} = \frac{x_{m,y}}{\sum_{m'=1}^{12} x_{m',y}}$, where $x_{m,y}$ is the number of cases reported during month $m$ in year $y$. As noted in a previous study in MA (8) and in another study in the US overall (9), the different age groups had a different pattern of seasonality. In children 0–5 yo and in adults ≥20 yo, higher proportions of cases were observed in July–

August; by contrast, in adolescents higher proportions of cases were observed during October–December, with a peak in November (Fig. S3). As a consequence, differences of synchrony were observed between age groups (Fig. S4): cases in infants 0–1 yo coincided with cases in 1–5 yo and adults, but preceded cases in adolescents by 3 mo.

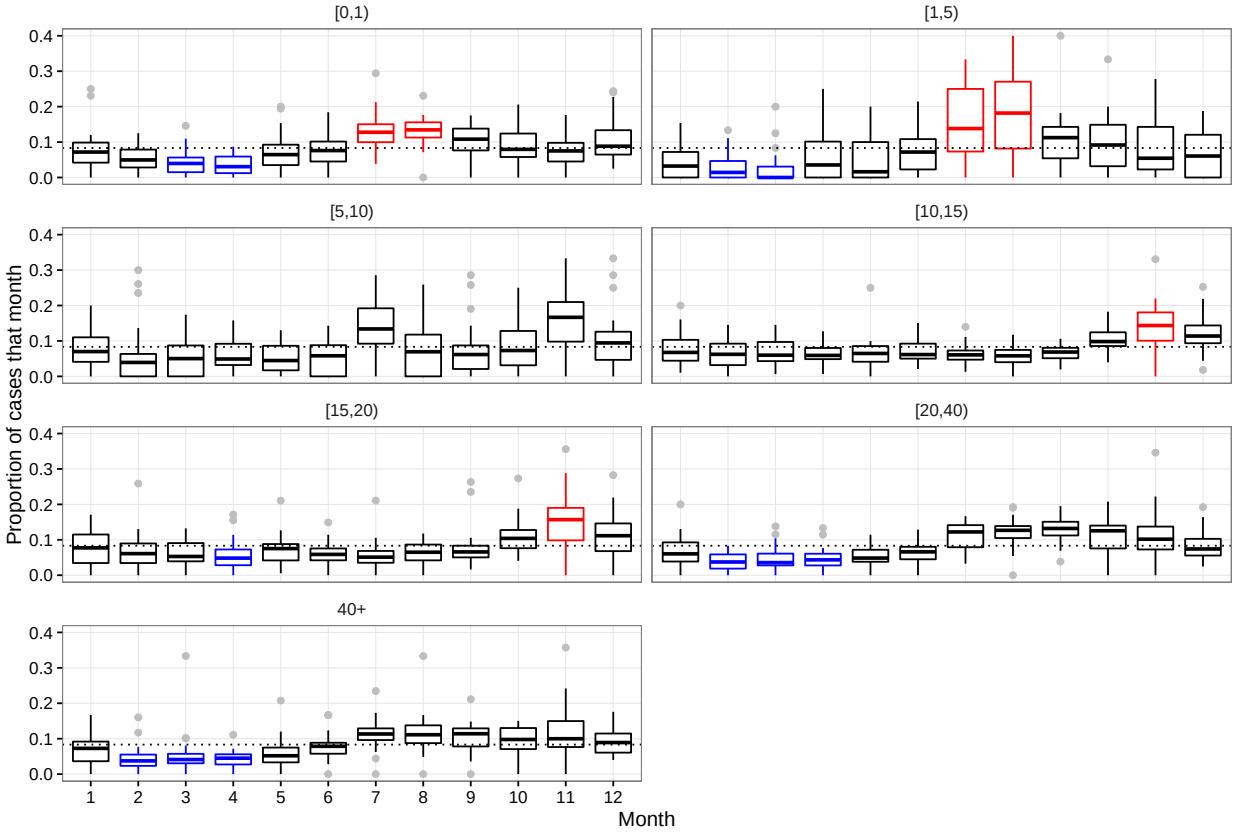


Figure S3: Age-specific seasonality in reported cases. The figure represents, for each age group, the year-to-year variability in the proportion of reported cases during month $m = 1, \ldots, 12$ , $\{p_{m,y}\}_{y=1,\ldots,n_y}$. **[Use color figures with bootstrap significance.]**

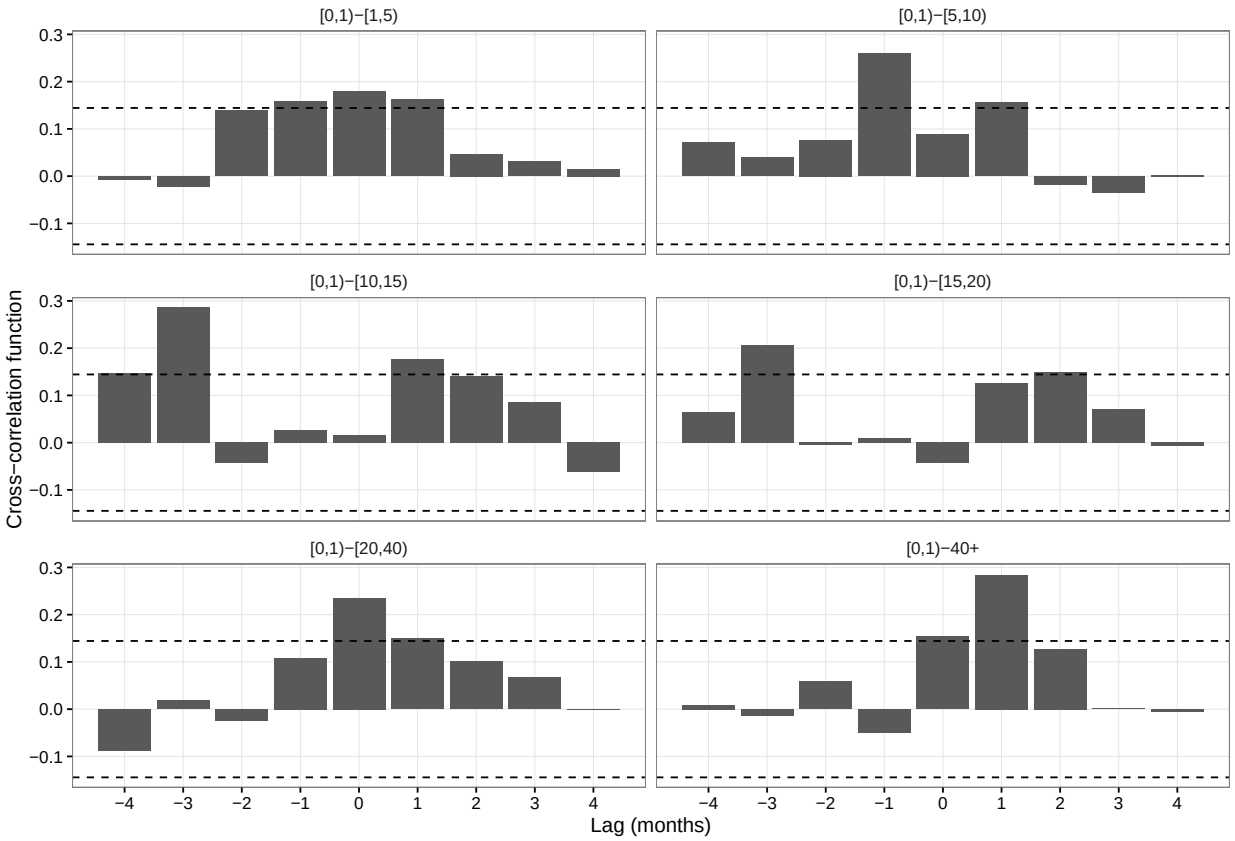


Figure S4: Cross-correlations between age groups, with age group 0–1 yo taken as the reference age group. Before computation, the data in each age group were prewhitened by taking the residuals from a seasonal ARIMA(2, 0, 0)(1, 0, 0)[12] model, identified as the most parsimonious model in infants 0–1 y of age (10).

## S1.4 Quantifying trends by age group

To quantify age-specific trends in reported cases, we fitted Poisson regression models for each age group, with time as covariate and population sizes as offsets:

$$\frac{x_t}{N_t} \sim \mathrm{P}(e^{\alpha+\beta t})$$

where $x_t$ represents the annual number of cases, $N_t$ the population size, and $t$ the time (in y). The trend was estimated as $100 \times (e^{\beta} - 1)$, which represents the annual relative variation (in percents) in the number of reported cases. These estimates, presented in Table S2, are reported in the main text.

| Age group | Trend (se) |
|---|---|
| [0,1) y | –0.9 (0.8) |
| [1,5) y | 2.0 (1.2) |
| [5,10) y | –2.7 (1.2) |
| [10,15) y | 6.2 (0.4) |
| [15,20) y | 9.3 (0.4) |
| [20,40) y | 17.0 (0.7) |
| 40+ y | 16.4 (0.7) |
| overall | 9.7 (0.2) |

Table S2: Age-specific trends (standard errors) estimated from the annual incidence data by Poisson regression.

## S1.5 Demographic data

Age-stratified mid-year population estimates in MA were available from the US census bureau for years 1990–2005 (11, 12). Annual number of births were shared by Kevin Bakker for years 1990–2005 (13). These demographic data, plotted in Fig. S5, were interpolated using smoothing splines (with 10 degrees of freedom) to calculate the time-varying annual number of births, $B(t)$; the age-stratified population sizes, $N_i(t)$; and the first derivative of the age-stratified population sizes, $\dot{N}_i(t)$. These quantities were used as covariates in the model so that the simulated population sizes approximately equalled the actual values (cf. Text S2.1.3).

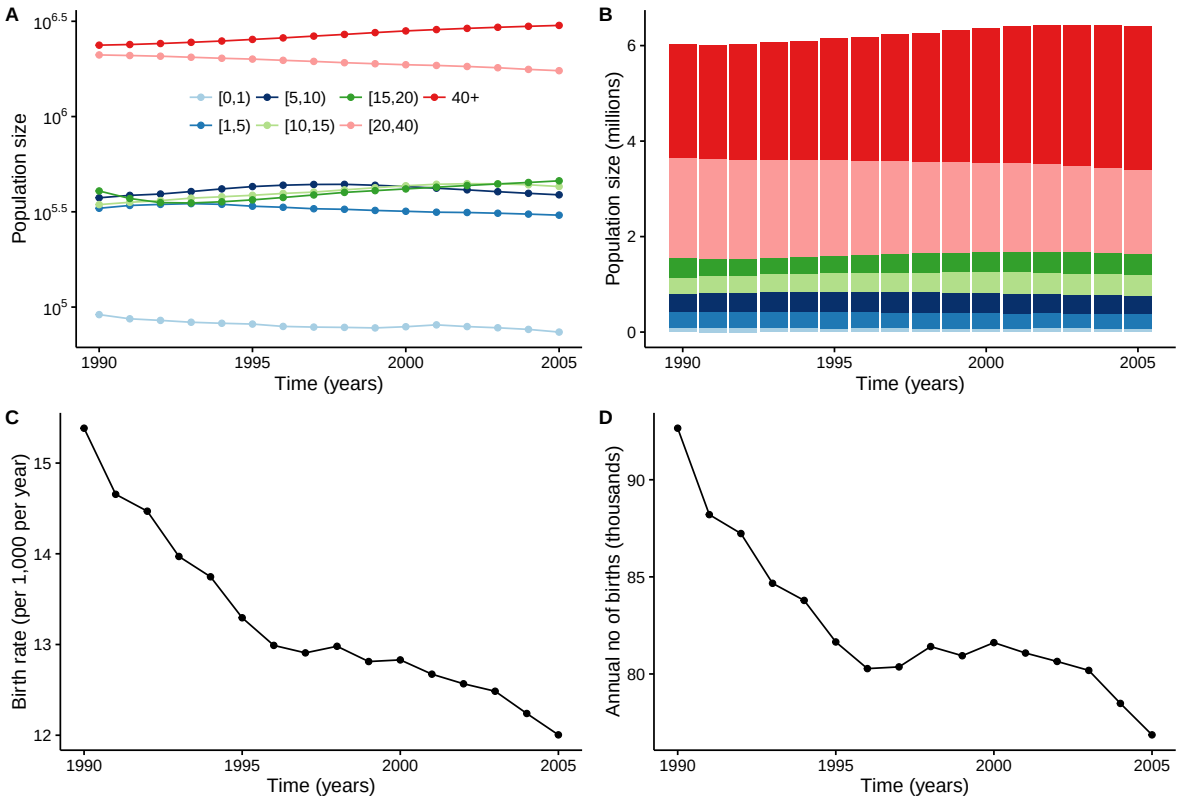


Figure S5: Demographic data, MA during 1990–2005. A, B: Age-specific population sizes; C: Annual birth rate (per 1000); D: Annual births (in thousands).

# S2 Supplementary methods

## S2.1 Model formulation

### S2.1.1 A model of pertussis transmission with different modes of loss of immunity

We implemented an age-stratified, compartmental model of pertussis transmission, building on previously described models (14–16). The model is an extension of the classic SEIR model that allows for post-vaccination infections in previously vaccinated or infected individuals. The population of susceptible is divided into those naive to exposure ($S^{(1)}$) and those whose immune system has been previously primed by vaccination or natural infection ($S^{(2)}$). Exposed and infected individuals are similarly divided into those who experience a naive infection ($E^{(1)}$ and $I^{(1)}$) or a post-vaccination infection ($E^{(2)}$ and $I^{(2)}$). Upon recovery from either type of infection, individuals move to the recovered class, $R$. To account for possible differences between infection- and vaccine-derived immunity, vaccinated individuals are explicitly modeled ($V$). Following recent advances in the literature (17–20), we considered 2 possible modes of failure of

infection- or vaccine-derived immunity:

1. Waning (failure in duration): immunized individuals lose their immunity and become susceptible ($S^{(2)}$) at a rate $\alpha_I$ for infection-derived immunity and $\alpha_V$ for vaccine-derived immunity.

2. Leakiness (failure in degree): immunized individuals ($R$ or $V$) remain susceptible to a post-vaccination infection ($E^{(2)}$), but at lower degree than susceptible individuals ($S^{(1)}$ or $S^{(2)}$). The degree of susceptibility is $\epsilon_I$ for infection-derived immunity and $\epsilon_V$ for vaccine-derived immunity.

For vaccine-derived immunity, we additionally considered a third mode of failure, for which, with probability $\epsilon_A$, vaccinated individuals immediately fail to mount an immune response and move the $S^{(1)}$ class (failure in take or primary vaccine failure (18)).

Individuals are categorized by 5-y age groups from age 5 y to age 75 y; the 0–5 y age group is further divided into 1–5 y and infants aged 0–4 mo and 4–12 mo. The 0–4 mo age group is included to represent the fact that infants are fully vulnerable to infection before receiving the second dose of DTP at age 4 months (16). Overall, the model consists of 17 age groups, labelled $i = 1, \ldots, 17$. Aging occurs continuously, at rates $\delta_i = \frac{1}{\Delta a_i}$ yr$^{-1}$, where $\Delta a_i$ is the age span in age group $i$. To model the effect of the primary vaccination course, a fraction $v_1$ of susceptible individuals ($S^{(1)}$ and $S_2^{(2)}$) is moved to the vaccinated class on aging from 0–4 mo to 4–12 mo. Similarly, the effect of the preschool booster dose is modeled by moving a fraction $v_2$ of susceptible individuals aging from 1–5 y to 5–10 y to the vaccinated class.

### S2.1.2 Contact network data

The model incorporated empirical age-specific contact rates from the Polymod study in Great Britain (21), corrected for reciprocity as detailed in the supplementary material of ref. (22). Let $C_{ij}$ be the average number of daily contacts (both physical and conversational, table S8.4.a in ref. (21)) reported by a participant of age group $i$ with members of age group $j$ (here individuals are categorized by 5-y age groups from age 0 to age 75 y, so that $1 \leq (i,j) \leq 15$). Denoting $N_i$ the number of individuals in age group $i$ in MA, the average total number of contacts between age groups $i$ and $j$ is: $E_{ij} = N_i C_{ij}$. Because of the necessary symmetry in the total number of contacts between age groups, the matrix $E = (E_{ij})$ was made symmetric: $E \to \frac{1}{2}(E + E^T)$. The individual average number of daily contacts between age groups $i$ and $j$, corrected for reciprocity, was then given by: $\forall (i,j), C_{ij} = \frac{E_{ij}}{N_i}$. The corrected matrix $C = (C_{ij})$ was used in all simulations and is plotted in Fig. S6.

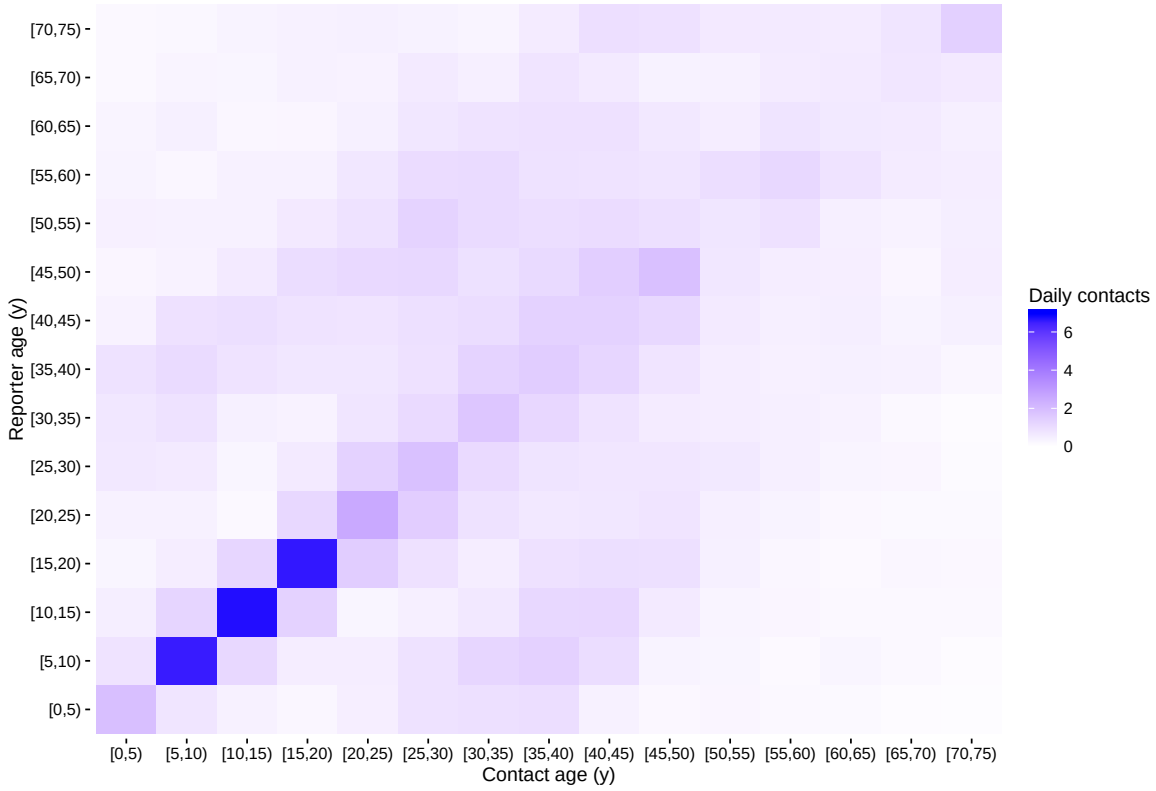


Figure S6: Age-specific contact matrix used in the simulations.

### S2.1.3 Model equations

The model equations for newborns (aged 0–4 mo, $i = 1$) are:

$$
\begin{aligned}
\frac{dV_1}{dt} &= 0 \\
\frac{dS_1^{(1)}}{dt} &= B(t) - (\lambda_1(t) + \delta_1 - \mu_1(t))S_1^{(1)} \\
\frac{dE_1^{(1)}}{dt} &= \lambda_1(t)S_1^{(1)} - (\sigma + \delta_1 - \mu_1(t))E_1^{(1)} \\
\frac{dI_1^{(1)}}{dt} &= \sigma E_1^{(1)} - (\gamma + \delta_1 - \mu_1(t))I_1^{(1)} \\
\frac{dS_1^{(2)}}{dt} &= \alpha_V V_1 + \alpha_I R_1 - (\lambda_1(t) + \delta_1 - \mu_1(t))S_1^{(2)} \\
\frac{dE_1^{(2)}}{dt} &= \lambda_1(t)(S_1^{(2)} + \epsilon_L V_1 + \epsilon_I R_1) - (\sigma + \delta_1 - \mu_1(t))E_1^{(2)} \\
\frac{dI_1^{(2)}}{dt} &= \sigma E_1^{(2)} - (\gamma + \delta_1 - \mu_1(t))I_1^{(2)} \\
\frac{dR_1}{dt} &= \gamma(I_1^{(1)} + I_1^{(2)}) - (\alpha_I + \epsilon_I\lambda_1 + \delta_1 - \mu_1(t))R_1
\end{aligned}
$$

Dynamics in older age groups ($i = 2, \ldots, 17$) are given by:

$$
\begin{aligned}
\frac{dV_i}{dt} &= \delta_{i-1}V_{i-1} + p_{i-1}(t)(1-\epsilon_A)\delta_{i-1}(S_{i-1}^{(1)} + c_{i-1}(t)S_{i-1}^{(2)}) - (\alpha_V + \epsilon_V\lambda_i(t) + \delta_i - \mu_i(t))V_i \\
\frac{dS_i^{(1)}}{dt} &= (1 - p_{i-1}(t)(1-\epsilon_A))\delta_{i-1}S_{i-1}^{(1)} - (\lambda_i(t) + \delta_i - \mu_i(t))S_i^{(1)} \\
\frac{dE_i^{(1)}}{dt} &= \delta_{i-1}E_{i-1}^{(1)} + \lambda_i(t)S_i^{(1)} - (\sigma + \delta_i - \mu_i(t))E_i^{(1)} \\
\frac{dI_i^{(1)}}{dt} &= \delta_{i-1}I_{i-1}^{(1)} + \sigma E_i^{(1)} - (\gamma + \delta_i - \mu_i(t))I_i^{(1)} \\
\frac{dS_i^{(2)}}{dt} &= (1 - p_{i-1}(t)(1-\epsilon_A)c_{i-1}(t))\delta_{i-1}S_{i-1}^{(2)} + \alpha_V V_i + \alpha_I R_i - (\lambda_i(t) + \delta_i - \mu_i(t))S_i^{(2)} \\
\frac{dE_i^{(2)}}{dt} &= \delta_{i-1}E_{i-1}^{(2)} + \lambda_i(t)(S_i^{(2)} + \epsilon_L V_i + \epsilon_I R_i) - (\sigma + \delta_i - \mu_i(t))E_i^{(2)} \\
\frac{dI_i^{(2)}}{dt} &= \delta_{i-1}I_{i-1}^{(2)} + \sigma E_i^{(2)} - (\gamma + \delta_i - \mu_i(t))I_i^{(2)} \\
\frac{dR_i}{dt} &= \delta_{i-1}R_{i-1} + \gamma(I_i^{(1)} + I_i^{(2)}) - (\alpha_I + \epsilon_I\lambda_i + \delta_i - \mu_i(t))R_i
\end{aligned}
$$

Here $\sigma$ represents the rate of progression from the exposed to the infectious class, $\gamma$ the recovery rate, $B(t)$ the time-varying number of births (cf section S1.5), and $p_i$ the age-specific vaccination coverage, such that: $\forall t, p_1(t) = v_1(t)$ (fraction of children having received 3 doses), $p_3(t) = v_2(t)$ (fraction of children having received a fifth dose), and $p_i(t) = 0$ for $i \notin \{1, 3\}$. The correction term $c_i(t) = \frac{\epsilon_A v_1(t)}{\epsilon_A v_1(t) + 1 - v_1(t)}\mathbb{1}(i = 3)$ is used to account for the fact that, on aging to age group 5–10 y, only children who received a primary course of vaccination (but for whom the vaccine did not take) can receive a fifth dose. The age-stratified migration rates, $\mu_i(t)$, were calculated from the smoothed demographic data (cf section S1.5) according to the equations: $\dot{N}_1(t) = B(t) + (\mu_1(t) - \delta_1)N_1(t)$ and $\dot{N}_i(t) = \delta_{i-1}N_{i-1}(t) + (\mu_i(t) - \delta_i)N_i(t)$.

The force of infection in age group $i = 1, \ldots, 17$ is given by:

$$\lambda_i = q_i \sum_{j=1}^{15} F_{ij}(t)\tilde{C_{ij}} \frac{\tilde{I}_j^{(1)} + \theta\tilde{I}_j^{(2)} + \iota}{\tilde{N}_j}$$

where $q_i$ is the probability of infection given exposure in age group $i$, $\theta$ the relative infectiousness of post-vaccination infections compared with naive infections, and $\iota$ an immigration term, fixed to $\iota = 1$ in all simulations. The total number of individuals in the $j$-th 5-y age block is denoted by $\tilde{N}_j$, such that $\tilde{N}_1 = N_1 + N_2 + N_3$ (age groups 0–4 mo, 4–12 mo, and 1–5 y) and $\forall j = 2, \ldots, 15, \tilde{N}_j = N_{j+2}$. Similarly, $\tilde{I}_1^{(k)} = I_1^{(k)} + I_2^{(k)} + I_3^{(k)}$ and $\forall j \geq 2, \tilde{I}_j^{(k)} = I_{j+2}^{(k)}$, where $k \in \{1, 2\}$. The matrix $\tilde{C}$ represents the "augmented" contact matrix, used to integrate the two extra age groups 0–4 mo and 4–12 mo. Because all individuals in the 0–5 y age group are assumed to have the same number of contacts, we have: $\forall j, \tilde{C_{1j}} = \tilde{C_{2j}} = \tilde{C_{3j}} = C_{1j}$

and $\forall i = 4, \ldots, 17, \forall j, \tilde{C_{ij}} = C_{i-2,j}$.

To capture the seasonality in children's contacts, we incorporated an age-dependent seasonal transmission term, $F_{ij}(t)$, modeled using $K = 3$ periodic spline basis functions with period 1 y, denoted by $\{S_k(t)\}_{k=1,\ldots,K}$. Because of the different seasonality in the number of reports between age groups 5–10 y and 10–20 y (Fig. S3), we allowed for different seasonality coefficients for these two age groups. For simplicity, seasonal forcing was also assumed assortative, and applied only within the same age group (i.e., only for contacts between 5–10 y and 5–10 y, between 10–15–y and 10–15 y, and between 15–20 y and 15–20 y). Therefore, we write $F_{42}(t) = \exp(\sum_{k=1}^{K} \omega_k^C S_k(t))$ for age group 5–10 y, $F_{53}(t) = F_{64}(t) = \exp(\sum_{k=1}^{K} \omega_k^T S_k(t))$ for age groups 10–15 y and 15–20 y, and $F_{ij}(t) = 1$ in all other cases. The function $F(t) = \exp(\sum_{k=1}^{K} \omega_k S_k(t))$ has geometric mean $\exp(\mathbb{E}_t(\log F(t))) = \exp(\frac{\sum_{k=1}^{K} w_k}{K})$. To ensure a geometric mean of 1 throughout the year, we therefore imposed the constraints $\sum_{k=1}^{K} \omega_k^C = 0$ and $\sum_{k=1}^{K} \omega_k^T = 0$ , so that only 2 seasonal coefficients were estimated, separately for age group 5–10 y and 10–20 y.

#### S2.1.4 Observation model

To complete the model specification, we model the observation process. Between times $t - \Delta t$ and $t$, where $\Delta t = 1$ month represents the reporting period, we denote by $C_{i,t}^{(1)}$ (resp. $C_{i,t}^{(2)}$) the complete number of new naive infections (resp. new post-vaccination infections) in age group $i$, counted as the number of transitions from $I_i^{(1)}$ (resp. $I_i^{(2)}$) to $R$ during that time period. The corresponding case report in age group $i$, $CR_{i,t}$, is modeled as a negative binomial: $CR_{i,t} \sim \text{NB}(\rho_i^{(1)}(C_{i,t}^{(1)} + \eta C_{i,t}^{(2)}), 1/\tau)$. Thus, $\mathbb{E}(CR_{i,t}|C_{i,t}^{(1)}, C_{i,t}^{(2)}) = \rho_i^{(1)}(C_{i,t}^{(1)} + \eta C_{i,t}^{(2)})$ and $\text{Var}(CR_{i,t}|C_{i,t}^{(1)}, C_{i,t}^{(2)}) = \rho_i^{(1)}(C_{i,t}^{(1)} + \eta C_{i,t}^{(2)}) + \tau(\rho_i^{(1)}(C_{i,t}^{(1)} + \eta C_{i,t}^{(2)}))^2$, where $\rho_i^{(1)}$ is the reporting probability of naive infections in age group $i$, $\eta$ the relative reporting probability of post-vaccination infections vs. naive infections (assumed age-independent), and $\tau$ the reporting overdispersion.

#### S2.1.5 Model variants and simulation protocol

We considered both deterministic and stochastic variants of the model. For the deterministic variant, the deterministic skeleton, described by the system of differential equations above, was integrated numerically. The stochastic variant was implemented as a continuous-time Markov process, approximated via a multinomial modification of the $\tau$-leap algorithm with a fixed time step $\Delta t = 10^{-2}$ yr (23). For this variant, to ensure that the simulations approximately maintained the correct population sizes in each age group $i$, all compartments, $X_i$, were updated deterministically via, $X_i \leftarrow [X_i e^{\mu_i(t)\Delta t}]$, where [.] is the operation of rounding to nearest integer value and $\mu_i(t)$ is the time-varying migration rate in age group $i$. For both variants, the simulations were started 70 y before the start of vaccination (that is, in 1870), assuming that the system was at equilibrium in the prevaccine era (16, 22).

### S2.1.6 Reproduction numbers and vaccine impact

Following ref. (18), we define $\varphi = \frac{1}{p}(1 - \frac{R_p}{R_0})$ a measure of vaccine impact, where $p$ is the vaccine coverage at birth, $R_0$ the basic reproduction number in the absence of vaccination ($p = 0$), and $R_p$ the basic reproduction number with vaccination at coverage $p$. To calculate these quantities, we considered a simplified model with 15 age groups (0–5 y, 5–10 y, ..., 70–75 y; $\delta_i = 1/5$ yr$^{-1}$ for $i = 1, \ldots, 15$), a constant birth rate $b = 1/75$ yr$^{-1}$, vaccination at birth at coverage $p = v_1$, no exposed class, no forcing in contact rates, and constant age-stratified population sizes (such that $\forall i, N_i = \frac{bN_s}{\delta_i}$ where $N_S$ is the total population size). The equations are

$$\begin{aligned}
\frac{dV_1}{dt} &= bNp(1-\epsilon_A) - (\alpha_V + \epsilon_V\lambda_1 + \delta_1)V_1 \\
\frac{dS_1^{(1)}}{dt} &= bN(1-p(1-\epsilon_A)) - (\lambda_1 + \delta_1)S_1^{(1)} \\
\frac{dI_1^{(1)}}{dt} &= \lambda_1 S_1^{(1)} - (\gamma + \delta_1)I_1^{(1)} \\
\frac{dS_1^{(2)}}{dt} &= \alpha_V V_1 + \alpha_I R_1 - (\lambda_1 + \delta_1)S_1^{(2)} \\
\frac{dI_1^{(2)}}{dt} &= \lambda_1(S_1^{(2)} + \epsilon_V V_1 + \epsilon_I R_1) - (\gamma + \delta_1)I_1^{(2)} \\
\frac{dR_1}{dt} &= \gamma(I_1^{(1)} + I_1^{(2)}) - (\alpha_I + \epsilon_I\lambda_1 + \delta_1)R_1
\end{aligned}$$

for age group 0–5 y and

$$\begin{aligned}
\frac{dV_i}{dt} &= \delta_{i-1}V_{i-1} - (\alpha_V + \epsilon_V\lambda_i + \delta_i)V_i \\
\frac{dS_i^{(1)}}{dt} &= \delta_{i-1}S_{i-1}^{(1)} - (\lambda_i + \delta_i)S_i^{(1)} \\
\frac{dI_i^{(1)}}{dt} &= \delta_{i-1}I_{i-1}^{(1)} + \lambda_i S_i^{(1)} - (\gamma + \delta_i)I_i^{(1)} \\
\frac{dS_i^{(2)}}{dt} &= \delta_{i-1}S_{i-1}^{(2)} + \alpha_V V_i + \alpha_I R_i - (\lambda_i + \delta_i)S_i^{(2)} \\
\frac{dI_i^{(2)}}{dt} &= \delta_{i-1}I_{i-1}^{(2)} + \lambda_i(S_i^{(2)} + \epsilon_V V_i + \epsilon_I R_i) - (\gamma + \delta_i)I_i^{(2)} \\
\frac{dR_i}{dt} &= \delta_{i-1}R_{i-1} + \gamma(I_i^{(1)} + I_i^{(2)}) - (\alpha_I + \epsilon_I\lambda_i + \delta_i)R_i
\end{aligned}$$

for older age age groups, where $\forall i = 1, \ldots, 15, \lambda_i = q_i \sum_{j=1}^{15} C_{ij} \frac{I_j^{(1)} + \theta I_j^{(2)}}{N_j}$.

We define, for $i = 1, \ldots, 15$, the vectors $\mathcal{F}_i^{(1)} = \lambda_i S_i^{(1)}$, $\mathcal{F}_i^{(2)} = \lambda_i(S_i^{(2)} + \epsilon_V V_i + \epsilon_I R_i)$, $\mathcal{V}_i^{(1)} = (\gamma + \delta_i)I_i^{(1)} -$

$\delta_{i-1} I_{i-1}^{(1)}$, $\mathcal{V}_i^{(2)} = (\gamma + \delta_i) I_i^{(2)} - \delta_{i-1} I_{i-1}^{(2)}$ (by convention $\delta_0 = 0$). We form the matrices $F^{11}$, $F^{12}$, $F^{21}$, and $F^{22}$, such that:

$$\begin{aligned}
F_{ij}^{11} &= \frac{\partial \mathcal{F}_i^{(1)}}{\partial I_j^{(1)}} &&= q_i \frac{c_{ij}}{N_j} S_i^{(1)} \\
F_{ij}^{12} &= \frac{\partial \mathcal{F}_i^{(1)}}{\partial I_j^{(2)}} &&= \theta F_{ij}^{11} \\
F_{ij}^{21} &= \frac{\partial \mathcal{F}_i^{(2)}}{\partial I_j^{(1)}} &&= q_i \frac{c_{ij}}{N_j} (S_i^{(2)} + \epsilon_V V_i + \epsilon_I R_i) \\
F_{ij}^{22} &= \frac{\partial \mathcal{F}_i^{(2)}}{\partial I_j^{(2)}} &&= \theta F_{ij}^{21}
\end{aligned}$$

We similarly form the matrices $V^{11}$, $V^{12}$, $V^{21}$, and $V^{22}$:

$$\begin{aligned}
V_{ij}^{11} &= \frac{\partial \mathcal{V}_i^{(1)}}{\partial I_j^{(1)}} &&= (\gamma + \delta_i) \mathbb{I}_{i,j} - \delta_{i-1} \mathbb{I}_{i-1,j} \\
V_{ij}^{12} &= \frac{\partial \mathcal{V}_i^{(1)}}{\partial I_j^{(2)}} &&= 0 \\
V_{ij}^{21} &= \frac{\partial \mathcal{V}_i^{(2)}}{\partial I_j^{(1)}} &&= 0 \\
V_{ij}^{22} &= \frac{\partial \mathcal{V}_i^{(2)}}{\partial I_j^{(2)}} &&= (\gamma + \delta_i) \mathbb{I}_{i,j} - \delta_{i-1} \mathbb{I}_{i-1,j}
\end{aligned}$$

where $\mathbb{I}_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ is the Kronecker delta. Defining $F = \begin{bmatrix} F^{11} & F^{12} \\ F^{21} & F^{22} \end{bmatrix}$ and $V = \begin{bmatrix} V^{11} & V^{12} \\ V^{21} & V^{22} \end{bmatrix}$, the next-generation matrix is given by $G = FV^{-1}$, calculated at the disease-free equilibrium, with or without vaccination. In the absence of vaccination ($p = 0$), the disease-free equilibria are given by $\forall i, V_i^* = S_i^{(2)^*} = R_i^* = 0$ and $S_i^{(1)^*} = N_i$ and we write $R_0 = \rho(FV^{-1})$ the leading eigenvalue of the next-generation matrix. With vaccination at coverage $p$, the disease-free equilibria are given by $\forall i, S_i^{(1)^*} = (1 - p(1 - \epsilon_A))N_i, V_i^* = p(1-\epsilon_A)N_i \prod_{k=1}^{i} \frac{\delta_k}{\delta_k + \alpha_V}, S_i^{(2)^*} = N_i - V_i^* - S_i^{(1)^*}, R_i^* = 0$; $R_p = \rho(FV^{-1})$. The two reproduction numbers $R_0$ and $R_p$ were calculated for each fitted model and are presented in the all the tables of parameter estimates below.

## S2.2 Model parametrization

### S2.2.1 Fixed parameters

The average latent period ($D_E = 8/365$ yr), and the average infectious period ($D_I = 15/365$ yr), were fixed at values used in previous models (16, 20, 22). For the deterministic model, these durations were converted to

annual rates, via $\sigma = 1/D_E$ and $\gamma = 1/D_I$. For the stochastic model, we used a discrete-time correction (23), such that $\sigma = -\frac{1}{\Delta t}\log(1-\frac{\Delta t}{D_E})$ and $\gamma = -\frac{1}{\Delta t}\log(1-\frac{\Delta t}{D_I})$. According to previous epidemiological studies (14–16), we assumed perfect infection-derived immunity, i.e., $\alpha_I = \epsilon_I = 0$. The vaccination coverages for the primary series and the preschool booster dose were fixed at $v_1 = 0.97$ and $v_2 = 0.93$. The smoothed annual number of births, $B(t)$, and the age-stratified annual migration rates, $\mu_i(t)$, were fixed and included as covariates in the model.

According to previous evidence (16), the probabilities of infection given exposure, $q_i$, were allowed to decrease with age. For simplicity, we estimated only three parameters in 0–10 y (parameter $q_1$), in 10–20 y (relative susceptibility to infection in 10–20 y to that in 0–10 y, parameter $\frac{q_2}{q_1}$), and in $\geq$ 20 y (relative susceptibility to infection in $\geq$ 20 y to that in 10–20 y, parameter $\frac{q_3}{q_2}$). Despite scarce evidence, the reporting fidelity of pertussis is believed to decrease with age (24, 25). Based on US data during 1985–1988, Sutter and Cochi evaluated the completeness of reported pertussis hospitalizations, indeed found to decrease with age (26). As an aside, these authors also calculated the complete number of pertussis cases during that period, by dividing their estimates of the age-stratified complete number of pertussis hospitalizations by the age-specific probabilities of hospitalization. Comparing this number to the total number of reported cases during that period, the completeness of reported cases was estimated at 11.6% (26). We repeated this simple calculation to estimate how the completeness of reported cases varies with age. Because the age-specific reported cases were not reported in the study by Sutter and Cochi, we used the age distribution of reported cases during 1980–1989 in the US as a proxy (27). Thus, we multiplied the total number of reported cases indicated by Sutter and Cochi (14,057) by the proportion of cases by age group indicated in ref. (27) to estimate the age-specific reported cases during 1985–1988. These figures were then divided by the age-specific true number of cases estimated by Sutter and Cochi to calculate age-specific reporting probabilities (Table S3). The estimated reporting probabilities were 0.15 (0.09–0.27) in 0–1 y, 0.09 (0.04, 0.09) in 1–5 y, and 0.04 (0.01, 0.04) in $\geq$ 5 y. We fixed the reporting probabilities of naive infections in 0–1 y, 1–5 y, 5–10 y according to these estimates. By contrast, because of the use of serological testing in $\geq$ 10 y, the reporting probability of naive infections could not be fixed and was estimated in these age groups (parameter $\rho_1(10+) = \rho^{(1)}_{i\geq 5}$), as well as the relative reporting probability of post-vaccination infections (parameter $\eta$, assumed independent of age). The values of the fixed parameters are indicated in Table S4.

| Age class | <1 y | 1–4 y | 5–9 y | ≥10 y | Source |
|---|---|---|---|---|---|
| True no of hospitalizations | 10311 | 2167 | 561 | 1229 | (26), Table 2 |
| Hospitalization probability | 0.235 (0.147–0.42) | 0.056 (0.023–0.058) | 0.017 (0.004–0.017) | 0.017 (0.004–0.017) | (26), Table 6 |
| True no of cases | 43877 (24550–70143) | 38696 (37362–94217) | 33000 (33000–140250) | 72294 (72294–307250) | Calculated |
| Total no cases reported | 14057 | | | | CDC, (26) |
| Proportion of cases | 0.471 | 0.252 | 0.09 | 0.187 | (27), Table 2 |
| No cases reported | 6621 | 3542 | 1265 | 2629 | Calculated |
| Reporting probability | 0.15 (0.09–0.27) | 0.09 (0.04–0.09) | 0.04 (0.01–0.04) | 0.04 (0.01–0.04) | Calculated |

Table S3: Estimates of age-specific reporting probabilities.

| Symbol | Parameter | Value |
|---|---|---|
| $D_E$ | Latent period | 8 days |
| $D_I$ | Infectious period | 15 days |
| $(C_{ij})$ | Age-specific contact rates | Fig. S6 |
| $\alpha_I$ | Waning rate of infection-derived immunity ($\mathrm{yr}^{-1}$) | 0 |
| $\epsilon_I$ | Leakiness of infection-derived immunity | 0 |
| $\rho_1^{(1)}, \rho_2^{(1)}$ | Reporting probability of naive infections in 0–1 y | 0.15 |
| $\rho_3^{(1)}$ | Reporting probability of naive infections in 1–5 y | $0.5 \times 0.15$ |
| $\rho_4^{(1)}$ | Reporting probability of naive infections in 5–10 y | $0.25 \times 0.15$ |
| $t_0$ | Start date of vaccination | 1940 |
| $t_1$ | End date of linear ramp-up for vaccination | 1955 |
| $\delta_{i=1,\ldots,17}$ | Aging rates ($\mathrm{yr}^{-1}$) | $(\frac{12}{4}, \frac{12}{8}, \frac{1}{4}, \frac{1}{5}, \ldots, \frac{1}{5})$ |

Table S4: Fixed model parameters.

### S2.2.2 Estimated parameters and estimation procedure

To determine the mode of vaccine-derived immunity, we used likelihood-based inference to evaluate the support of three models:

1. **No loss of vaccine-derived immunity (no-loss model).** After an initial failure in take, vaccine-derived is hypothesized to be perfect, so that no post-vaccine infections are possible. For this model, the estimated parameters were $\epsilon_A$ (fraction of primary vaccine failures), $q_1$, $q_2/q_1$, $q_3/q_2$ (susceptibility factors), $\omega_{1,2}^C$ (seasonality coefficients in children 5–10 y), $\omega_{1,2}^T$ (seasonality coefficients in adolescents 10–20 y), $\rho_1(10+)$ (reporting probability of naive infections in $\geq$ 10 y), and $\tau$ (reporting overdispersion).

2. **Waning vaccine-derived immunity (waning model).** After an initial failure in take, vaccine-derived immunity is hypothesized to wane at rate $\alpha_V$. The estimated parameters are those of the no-loss model, plus $\alpha_V$, $\theta$ (relative transmissibility of post-vaccine infections to that of naive infections), and $\eta$ (relative reporting probability of post-vaccine infections to that of naive infections).

3. **Leaky vaccine-derived immunity (leaky model).** After an initial failure in take, vaccine-derived

immunity is hypothesized to be leaky, with degree of leakiness $\epsilon_V$. The estimated parameters are those of the no-loss model, plus $\epsilon_V$, $\theta$ (relative transmissibility of post-vaccine infections to that of naive infections), and $\eta$ (relative reporting probability of post-vaccine infections to that of naive infections).

For each model, parameters were estimated in two steps:

1. **Trajectory matching**. The deterministic variant of the model was fitted to the data using maximum likelihood estimation via trajectory matching. In this case, the observation model is the only source of variability in simulated observations, and the likelihood can be calculated exactly. The likelihood was maximized using the subplex algorithm, implemented in the R package `nloptr` (28). The search was initiated over $10^4$ starting points generated using Latin hypercube sampling over broad parameter ranges (Table S5). To ensure convergence to the maximum likelihood estimate (MLE), the optimization was repeated on the 500 best parameter sets. A parametric bootstrap was then used to assess uncertainty in parameter estimates. For each model, 500 synthetic time series of simulated data were generated at the MLE. For each of these 500 synthetic data sets, parameters were re-estimated as described above, resulting in a bootstrap distribution of parameter estimates. 95% confidence intervals for the estimated parameters and for the derived parameters $R_0$ (basic reproduction number), $R_p$ (vaccine reproduction number), $\varphi$ (vaccine impact), and $\rho_2(10+) = \eta\rho_1(10+)$ (reporting probability of post-vaccine infections in $\geq 10\,\mathrm{y}$) were then calculated from the bootstrap distribution.

2. **Maximum iterated filtering.** The stochastic variant of the model was fitted using the maximum iterated filtering (MIF) algorithm (29), implemented in the R `pomp` package version 1.2.1.1 (30). The following algorithmic parameters were used: 2000 particles; 50 MIF iterations; random walk intensity of $10^{-6}$ during the first MIF iteration and $10^{-2}$ for the next 49 iterations. Because the model was simulated for a long period before the first data point (January 1990), a time-varying random walk was used for the parameters, with no perturbation until the first data point. For each MIF run, the log-likelihood was computed as the log of the average likelihood of 10 replicate particle filters, each with $5 \times 10^3$ particles; the standard error of the log-likelihood estimate was computed from these replicates using a jackknife implemented in the function `logmeanexp` in the `pomp` package. Because each MIF run required ≈24 hours of computation, we sought to find good starting parameter values to initiate the algorithm. To do this, we calculated for each model the range of starting parameters of trajectory matching such that $\{\Theta : 2 \times \log \frac{L(\hat{\Theta})}{L(\Theta)} < \chi^2_{\dim(\Theta),0.95}\}$, where $\hat{\Theta}$ is the maximum likelihood estimate from trajectory matching, $\dim(\Theta)$ the number of estimated parameters, and $\chi^2_{\dim(\Theta),0.95}$ the 95% quantile of a $\chi^2$ distribution with $\dim(\Theta)$ degrees of freedom. The search was then initiated over 100 starting points generated using Latin hypercube sampling over that range. As for trajectory

matching, the estimations were repeated from the best parameter sets to ensure convergence to the MLE and a parametric bootstrap was used to generate a bootstrap distribution of size 100.

| Symbol | Parameter | Start range |
|---|---|---|
| $q_1$ | Susceptibility factor in 0–10 y | [0, 1] |
| $\frac{q_2}{q_1}$ | Relative susceptibility in 10–20 y vs. 0–10 y | [0, 1] |
| $\frac{q_3}{q_2}$ | Relative susceptibility in $\geq$ 20 y vs. 10–20 y | [0, 1] |
| $\omega_C^1, \omega_C^2, \omega_T^1, \omega_T^2$ | Seasonality coefficients | $[-5, 5]$ |
| $\theta$ | Relative infectiousness of post-vaccine infections | [0, 1] |
| $\epsilon_A$ | Probability of primary vaccine failure | [0, 1] |
| $\alpha_V$ | Waning rate of vaccine-derived immunity | [0, 10] $\text{yr}^{-1}$ |
| $\epsilon_V$ | Leakiness of vaccine-derived immunity | [0, 1] |
| $\rho_1(10+)$ | Reporting probability of naive infections in 10–20 y | [0, 1] |
| $\eta$ | Relative reporting probability of repeat infections | [0, 1] |
| $\tau$ | Reporting overdispersion | [0, 10] |

Table S5: Parameters ranges used to generate starting parameter sets for trajectory matching.

# S3 Supplementary results

## S3.1 Parameter estimates

The parameter estimates of the base model (wP- and aP-derived immunity assumed identical, perfect infection-derived immunity) are presented in Table S8 (deterministic variant) and in Table S7 (stochastic variant). In Fig. S7, we plot the estimated seasonal transmission term, given by $\exp(\sum_{k=1}^{3} \hat{\omega}_k^C S_k(t))$ in 5–10 y and by $\exp(\sum_{k=1}^{3} \hat{\omega}_k^T S_k(t))$ in 10–20 y (cf. Text S2.1.3). Notably, this term differed in the two age groups: it peaked during April–May in 5–10 y, and later in the year (August–September) in 10–20 y. Although we did not investigate further, we note that this result might explain the phase delay in reported cases observed in 10–20 y (Fig. S4).

| Quantity | No-loss model | Leaky model | Waning model |
|---|---|---|---|
| $\log L$ | –3726.9 (se: 0.4) | –3664.9 (se: 0.5) | –3594.5 (se: 0.6) |
| AIC | 7474 | 7356 | **7215** |
| $\Delta$AIC | 259 | 141 | 0 |
| $R_p$ | 2.4 (1.8, 2.7) | 1.6 (1.3, 2.2) | 1.8 (1.5, 2.0) |
| $R_0$ | 13.6 (7.5, 23.0) | 12.6 (9.0, 19.4) | 10.1 (6.5, 17.2) |
| Vaccine impact | 0.85 (0.70, 0.95) | 0.85 (0.75, 0.93) | 0.90 (0.81, 0.95) |

Table S6: **Model comparison**. The results are presented for the stochastic variant of the model, estimated using the MIF algorithm. The best AIC value is indicated in boldface.

| Quantity | No-loss model | Waning model | Leaky model |
|---|---|---|---|
| $\log L$ | –3726.9 (se: 0.4) | –3594.5 (se: 0.6) | –3664.9 (se: 0.5) |
| AIC | 7474 | **7215** | 7356 |
| $\Delta$AIC | 255 | 0 | 135 |
| $q_1$ | 0.10 (0.05, 0.16) | 0.09 (0.06, 0.15) | 0.09 (0.07, 0.14) |
| $q_2$ | 0.09 (0.04, 0.17) | 0.05 (0.03, 0.11) | 0.09 (0.06, 0.13) |
| $q_3$ | 0.032 (0.014, 0.063) | 0.008 (0.005, 0.021) | 0.030 (0.016, 0.044) |
| $\rho_1$(10+) | 0.89 (0.42, 0.98) | 0.63 (0.28, 1.00) | 0.60 (0.49, 1.00) |
| $\eta$ | — | 0.39 (0.19, 1.00) | 0.42 (0.16, 0.75) |
| $\rho_2$(10+) | — | 0.24 (0.10, 0.66) | 0.25 (0.12, 0.57) |
| $\theta$ | — | 0.99 (0.40, 1.00) | 0.71 (0.20, 1.00) |
| $\tau$ | 0.28 (0.24, 0.35) | 0.19 (0.14, 0.24) | 0.31 (0.26, 0.36) |
| $\epsilon_A$ | 0.15 (0.05, 0.27) | 0.04 (0.01, 0.08) | 0.05 (0.03, 0.17) |
| $\alpha_V$ ($\mathrm{yr}^{-1}$) | — | 0.011 (0.003, 0.021) | — |
| $\epsilon_V$ | — | — | 0.06 (0.02, 0.14) |
| $R_0$ | 13.6 (7.5, 23.0) | 10.1 (6.5, 17.2) | 12.6 (9.0, 19.4) |
| $R_p$ | 2.4 (1.8, 2.7) | 1.8 (1.5, 2.0) | 1.6 (1.3, 2.2) |
| $\varphi$ | 0.85 (0.70, 0.95) | 0.85 (0.75, 0.93) | 0.90 (0.81, 0.95) |

Table S7: **Parameter estimates of the stochastic variant of the base model (similar aP- and wP-derived immunity, perfect infection-derived immunity)**. For each model tested (no loss of vaccine-derived immunity, waning vaccine-derived immunity, or leaky vaccine-derived), the maximum likelihood estimates (95% CI) from the MIF algorithm are presented. The standard error (se) of the log-likelihood was calculated using 20 replicate particle filters, each with $5 \times 10^4$ particles. The best AIC value is indicated in boldface. Please refer to Table S5 for the meaning of the parameters.

| Quantity | No-loss model | Waning model | Leaky model |
|---|---|---|---|
| $\log L$ | –3915.2 | –3787.9 | –3873.7 |
| AIC | 7850 | **7602** | 7773 |
| $\Delta$AIC | 631 | 383 | 554 |
| $q_1$ | 0.07 (0.06, 0.09) | 0.07 (0.05, 0.10) | 0.09 (0.07, 0.13) |
| $q_2$ | 0.07 (0.06, 0.09) | 0.03 (0.02, 0.07) | 0.09 (0.06, 0.12) |
| $q_3$ | 0.022 (0.016, 0.026) | 0.005 (0.004, 0.014) | 0.029 (0.017, 0.041) |
| $\rho_1$(10+) | 0.83 (0.70, 1.00) | 0.89 (0.23, 1.00) | 0.53 (0.45, 1.00) |
| $\eta$ | — | 0.28 (0.15, 1.00) | 0.65 (0.61, 0.74) |
| $\rho_2$(10+) | — | 0.25 (0.08, 0.53) | 0.34 (0.09, 0.78) |
| $\theta$ | — | 1.00 (0.46, 1.00) | 1.00 (0.21, 1.00) |
| $\tau$ | 0.73 (0.65, 0.80) | 0.56 (0.50, 0.62) | 0.69 (0.61, 0.74) |
| $\epsilon_A$ | 0.21 (0.17, 0.25) | 0.05 (0.02, 0.07) | 0.04 (0.02, 0.06) |
| $\alpha_V$ ($\mathrm{yr}^{-1}$) | — | 0.018 (0.007, 0.044) | — |
| $\epsilon_V$ | — | — | 0.04 (0.03, 0.17) |
| $R_0$ | 10.4 (8.4, 12.5) | 7.6 (5.7, 11.6) | 12.8 (9.2, 17.1) |
| $R_p$ | 2.4 (2.2, 2.5) | 1.7 (1.6, 1.8) | 1.4 (1.1, 1.6) |
| $\varphi$ | 0.79 (0.75, 0.83) | 0.79 (0.72, 0.88) | 0.92 (0.88, 0.95) |

Table S8: **Parameter estimates of the deterministic variant of the base model (similar aP- and wP-derived immunity, perfect infection-derived immunity)**. For each model tested (no loss of vaccine-derived immunity, waning vaccine-derived immunity, or leaky vaccine-derived), the maximum likelihood estimates (95% CI) of trajectory matching are presented. The best AIC value is indicated in boldface. Please refer to Table S5 for the meaning of the parameters.

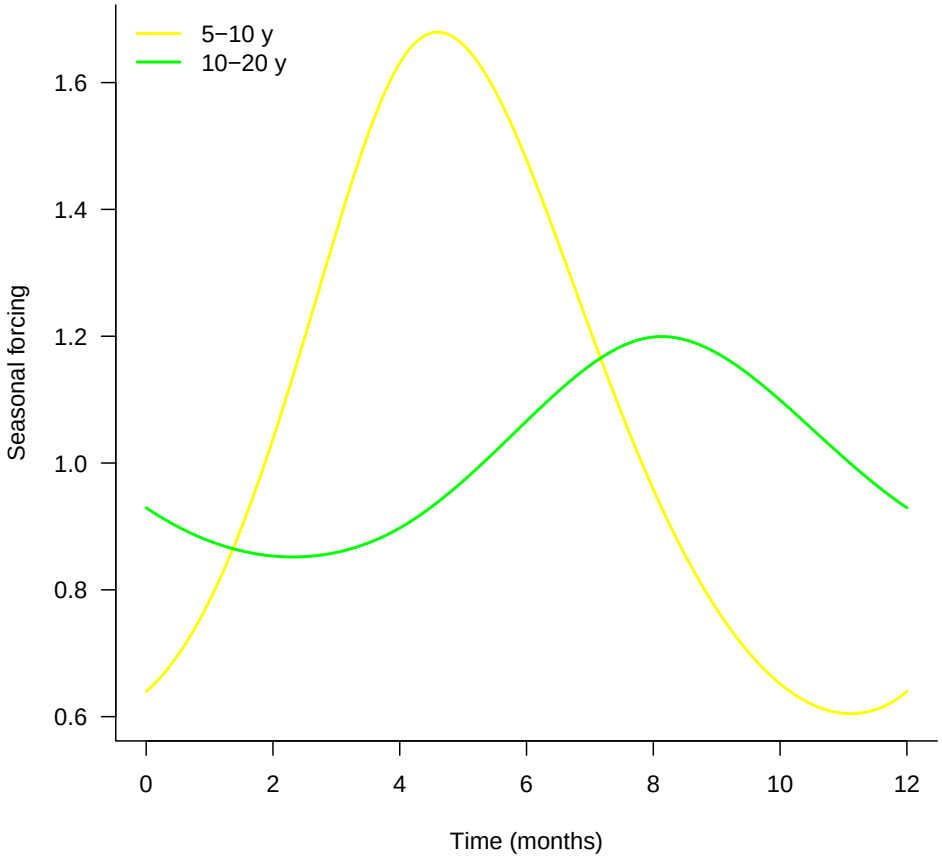


Figure S7: Estimated age-specific seasonal forcing.

## S3.2 Model assessment

We first assessed the fitted model by visual inspection of model simulations from 1990 (Fig. 2A in the main text). To initialize these simulations, we first ran particles from 1870 to the first data point (in January 1990) and calculated their log-likelihood at that point. We then used the particles most consistent with that data point as initial conditions for subsequent simulations.

To further quantify the agreement between model and data, we computed a generalized $R^2$, as defined in ref. (16). Let $d_{a,t}$ represent the observed reported incidence in age group $a$ during time $t$ and $m_{a,t}$ the corresponding simulated quantity. The generalized $R^2$ is then given by

$$R^2 = 1 - \frac{\sum_{a,t}[\log_{10}(1+d_{a,t}) - \log_{10}(1+m_{a,t})]^2}{\sum_{a,t}[\log_{10}(1+d_{a,t}) - \log_{10}(\bar{1}+d_a)]^2}$$

where $\log(\bar{1 + d_a}) = \frac{1}{T}\sum_t \log(1 + d_{a,t})$ is the average log-incidence in age group $a$. In this formula, the numerator represents the residual sum-of-squares of the fitted model and the denominator that of a null model with constant log-incidence (equal to the average log-incidence) by age group. Thus, $R^2$ measures the relative decrease ($R^2 > 0$), or increase ($R^2 < 0$), in the residuals sum-of-squares of our model compared with the null model. It quantifies the proportion of variance explained by the model to that not explained by age alone.

We calculated the $R^2$ for forecast horizons of 1, 6, and 12 months. To to this, we first divided the time period 1990–2005 into non-overlapping time blocks of 1, 6, or 12 months. For each time bock, we then initialized the state variables with their filtering means (calculated using a particle filter with $5 \times 10^3$ particles) one month before the beginning of the time block. Finally, we ran $5 \times 10^3$ model simulations over the time block and averaged across the simulations to calculate $m_{a,t}$. One-month-ahead predictions are presented in Fig. 2B in the main text; 6-mo- and 12-mo- ahead are presented in Fig. S8. The corresponding $R^2$ were 0.19 and 0.04.

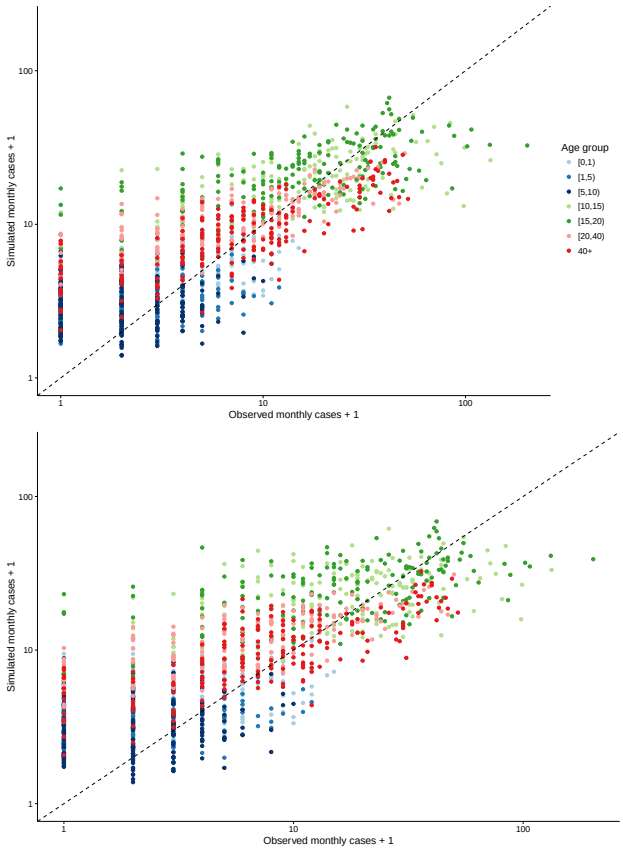


Figure S8: Comparison of model-data agreement at 6-mo and 12-mo forecast horizons.

### S3.3 Predictability as a function of the generation time

We conducted a simulation study to determine a baseline of dynamics predictability of our model for different generation times (7, 14, 21, or 28 days). We hypothesized that higher generation times would lead to lower predictability, as a result of less definite periodicity and noisier dynamics (31). For simplicity, we ignored the latent period ($\sigma \to \infty$), so that the infectious period equalled the generation time. For each generation time, we first back-calculated the susceptibility in 0–10 y (parameter $q_1$) to maintain a $R_0$ of 10 across simulations. We then generated 100 synthetic data sets of age-specific reported cases, assuming no reporting overdispersion ($\tau = 0$). Using each simulated data set as synthetic data, we calculated the generalized $R^2$ based on one-month-ahead predictions, as defined in the previous section.

As shown in the main text (Fig. 2C), the generalized $R^2$ indeed decreased with the generation time, while its variability increased. Notably, the $R^2$ remained below 60% for the shortest generation time, suggesting an intrinsic limit to predictability in complex (i.e., stochastic, nonlinear, and forced) age-structured models. In Fig. S10, we replicate Fig. 2B for a randomly chosen synthetic data set. Although the true model is known, the overestimation of troughs is evident. This is most readily explained by the fact that prediction means represent average simulations.

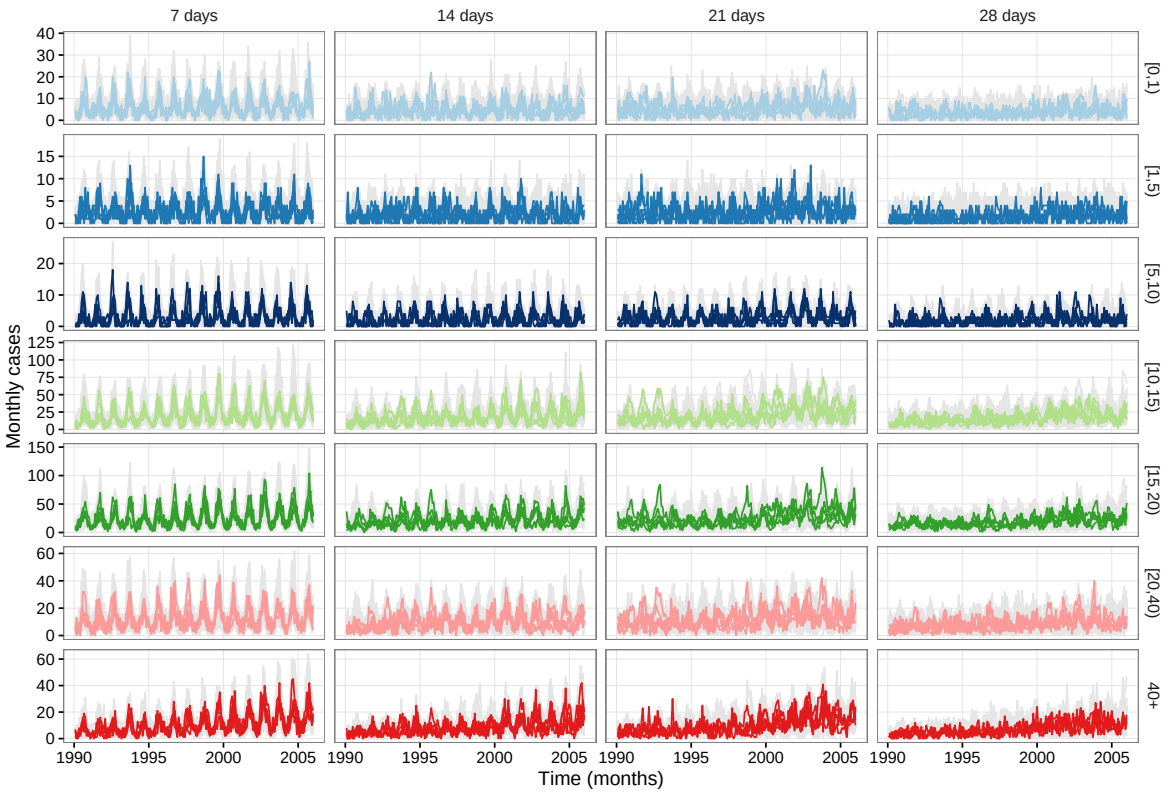


Figure S9: Data sets of monthly reports generated for the simulation study. Five stochastic realizations are highlighted.

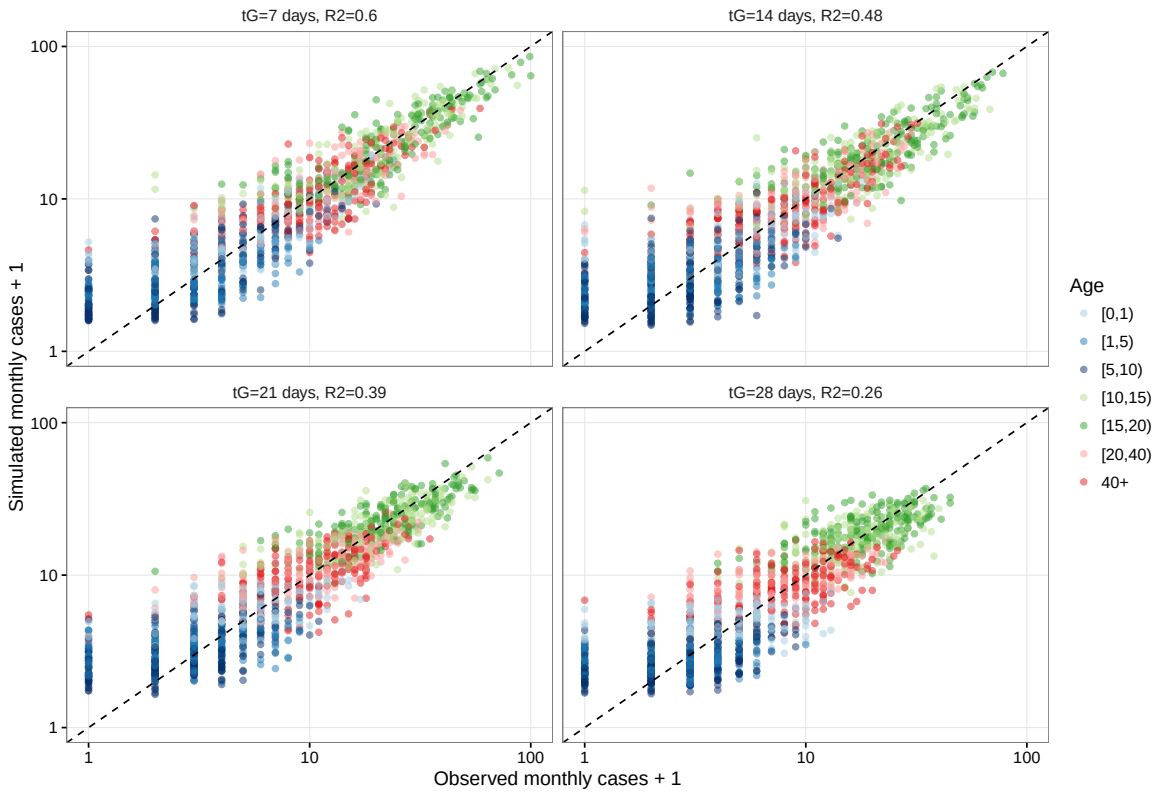


Figure S10: Quantitative comparison of model-data agreement for different generation times.

## S3.4 Sensitivity analyses

### S3.4.1 Assessing differences of immunity between aP and wP vaccines

To examine possible differences of immunity induced by wP and aP vaccines, we fitted an alternative model in which infection- and wP-derived immunity were assumed identical (with protected individuals in the $R$ compartment), but possibly different from aP-derived immunity (with protected individuals in the $V$ compartment). According to previous studies, we assumed that infection/wP-derived immunity was waning with average duration of protection 75 y (15). During the vaccine transition period, we also assumed, for simplicity, that the first dose of vaccine received determined the nature of subsequent immunity (32). Thus, all infants vaccinated before October 1996 were assumed to be protected by wP, while those vaccinated after were assumed to be protected by aP. We then repeated the estimations to determine which mechanism of loss of immunity (here specific to aP) was best supported by the data (Table S9).

| Quantity | Waning model | | Leaky model | |
|---|---|---|---|---|
| | TM | MIF | TM | MIF |
| $\log L$ | –3779.8 | –3611.2 (se: 1.3) | –3807.5 | –3622.8 (se: 0.4) |
| AIC | 7586 | **7248** | 7641 | 7272 |
| $\Delta$AIC | 338 | 0 | 393 | 24 |
| $q_1$ | 0.09 (0.06, 0.15) | 0.08 (0.05, 0.10) | 0.12 (0.08, 0.17) | 0.08 (0.06, 0.11) |
| $q_2$ | 0.04 (0.03, 0.09) | 0.04 (0.03, 0.05) | 0.06 (0.04, 0.12) | 0.04 (0.03, 0.06) |
| $q_3$ | 0.004 (0.003, 0.011) | 0.004 (0.003, 0.006) | 0.006 (0.005, 0.016) | 0.004 (0.003, 0.007) |
| $\rho_1(10+)$ | 1.00 (0.28, 1.00) | 0.89 (0.64, 0.99) | 1.00 (0.41, 1.00) | 0.67 (0.30, 1.00) |
| $\eta$ | 0.27 (0.13, 0.45) | 0.30 (0.14, 0.62) | 0.26 (0.12, 0.42) | 0.49 (0.20, 0.85) |
| $\rho_2(10+)$ | 0.27 (0.07, 0.41) | 0.27 (0.11, 0.56) | 0.26 (0.08, 0.38) | 0.33 (0.12, 0.68) |
| $\theta$ | 1.00 (0.24, 1.00) | 0.99 (0.50, 1.00) | 0.48 (0.13, 0.76) | 0.92 (0.55, 1.00) |
| $\tau$ | 0.55 (0.49, 0.60) | 0.22 (0.18, 0.33) | 0.58 (0.51, 0.63) | 0.19 (0.17, 0.26) |
| $\epsilon_A$ | 0.03 (0.01, 0.06) | 0.05 (0.02, 0.14) | 0.04 (0.01, 0.06) | 0.06 (0.02, 0.14) |
| $\alpha_V$ ($\mathrm{yr}^{-1}$) | 0.020 (0.019, 0.036) | 0.017 (0.002, 0.039) | — | — |
| $\epsilon_V$ | — | — | 0.07 (0.06, 0.13) | 0.05 (0.01, 0.07) |
| $R_0$ | 9.1 (7.1, 16.4) | 8.7 (5.4, 10.3) | 12.5 (9.0, 19.9) | 8.5 (6.5, 12.0) |
| $R_p$ | 2.1 (2.0, 2.6) | 1.9 (1.0, 2.4) | 1.2 (1.0, 1.5) | 1.1 (0.8, 1.5) |
| $\varphi$ | 0.79 (0.67, 0.90) | 0.80 (0.55, 0.93) | 0.93 (0.91, 0.98) | 0.89 (0.81, 0.95) |

Table S9: **Parameter estimates of the model with identical infection- and wP-derived immunity, but separate aP-derived immunity**. For each model tested (waning or leaky aP-derived immunity), the maximum likelihood estimates (95% CI) of trajectory matching (TM) and of the maximum iterated filtering algorithm (MIF) are presented. The best AIC value is indicated in boldface. Please refer to Table S5 for the meaning of the parameters.

### S3.4.2 Assessing the robustness of results to another contact matrix

In the absence of empirical contact data in the US, the results presented in the main text were obtained using the Polymod contact matrix in Great Britain. To assess the robustness of our results to this critical assumption, we calculated a contact matrix in MA using the method described in ref. (33). Briefly, the method uses highly detailed census and demographic data to build a matrix of "effective" contacts $M = (M_{ij})$. This matrix defines contacts among age groups up to a scaling constant $N_{tot}$, usually absorbed in the transmission rate (33). In our simulations, we fixed this constant to the total contact rate in the Polymod matrix. As shown in Fig. S11, the matrix presented fewer inter-generational contacts between children and adults, but more contacts between adults than the Polymod matrix. Repeating the estimations using that matrix, we found our main results to be robust (Table S10). Indeed, the waning model was preferred to the leaky model ($\Delta$AIC=153), with estimates comparable to those obtained with the Polymod contact matrix (waning rate 0.007 [0.001, 0.020] $\mathrm{y}^{-1}$, primary vaccine failure 0.03 [0.01, 0.05], vaccine impact 0.90 [0.80, 0.96]). We note, however, that the waning model was less consistent with the data using that matrix ($\Delta \log L = -41.4$, cf. Tables S7–S8).

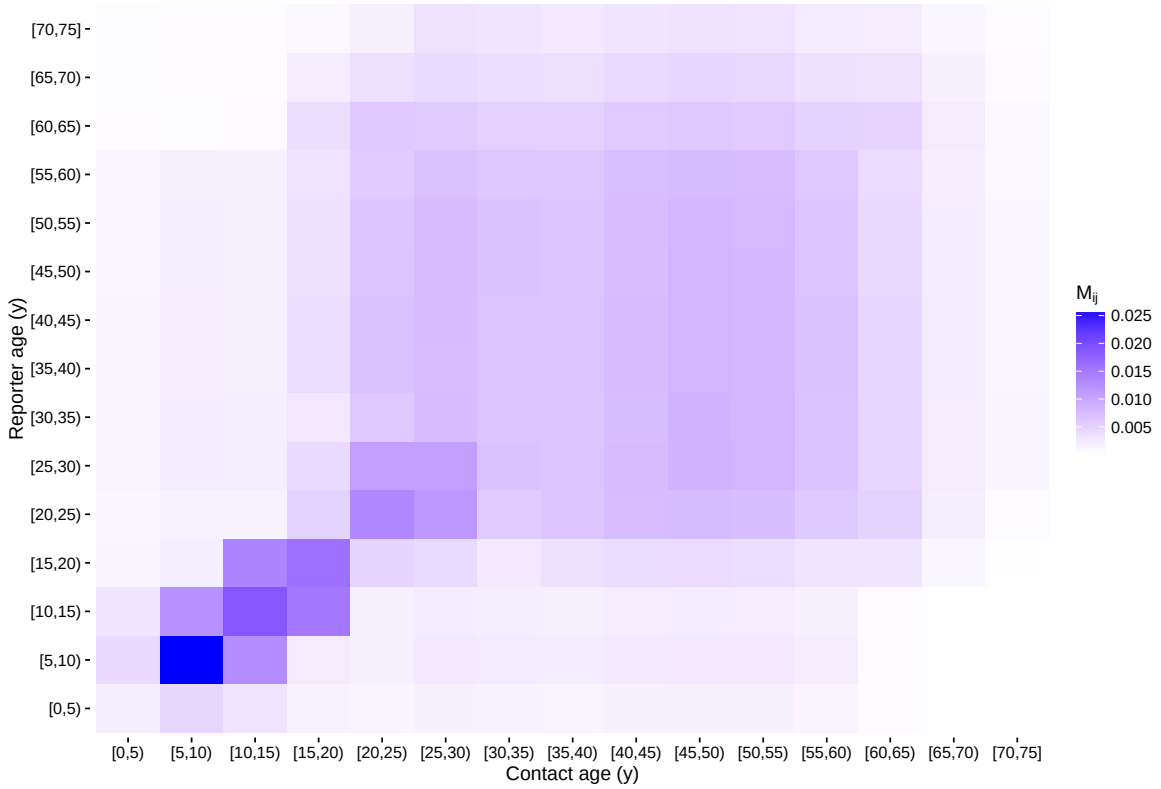


Figure S11: Contact matrix in MA.

| Quantity | Waning model | | Leaky model | |
|---|---|---|---|---|
| | TM | MIF | TM | MIF |
| $\log L$ | –3795.9 | –3637.8 (se: 0.5) | –3884.3 | –3714.5 (se: 0.4) |
| AIC | 7618 | **7302** | 7795 | 7455 |
| $\Delta$AIC | 316 | 0 | 493 | 153 |
| $q_1$ | 0.09 (0.05, 0.14) | 0.18 (0.10, 0.30) | 0.13 (0.09, 0.23) | 0.17 (0.12, 0.29) |
| $q_2$ | 0.05 (0.04, 0.10) | 0.09 (0.05, 0.24) | 0.13 (0.09, 0.22) | 0.13 (0.09, 0.27) |
| $q_3$ | 0.008 (0.006, 0.017) | 0.014 (0.007, 0.058) | 0.039 (0.025, 0.072) | 0.042 (0.027, 0.098) |
| $\rho_1(10+)$ | 0.54 (0.26, 0.87) | 0.74 (0.41, 0.98) | 0.64 (0.53, 1.00) | 0.81 (0.64, 1.00) |
| $\eta$ | 0.30 (0.16, 0.82) | 0.48 (0.06, 0.94) | 0.35 (0.08, 1.00) | 0.27 (0.01, 0.88) |
| $\rho_2(10+)$ | 0.16 (0.06, 0.36) | 0.36 (0.06, 0.77) | 0.22 (0.06, 0.97) | 0.22 (0.01, 0.80) |
| $\theta$ | 1.00 (0.54, 1.00) | 0.89 (0.23, 1.00) | 1.00 (0.08, 1.00) | 0.51 (0.05, 1.00) |
| $\tau$ | 0.58 (0.51, 0.64) | 0.28 (0.23, 0.33) | 0.71 (0.64, 0.77) | 0.32 (0.27, 0.38) |
| $\epsilon_A$ | 0.04 (0.02, 0.07) | 0.03 (0.01, 0.05) | 0.05 (0.03, 0.08) | 0.05 (0.01, 0.16) |
| $\alpha_V$ ($\mathrm{yr}^{-1}$) | 0.016 (0.007, 0.034) | 0.007 (0.001, 0.020) | 0 | 0 |
| $\epsilon_V$ | 0 | 0 | 0.04 (0.01, 0.11) | 0.05 (0.01, 0.16) |
| $R_0$ | 7.5 (5.1, 12.8) | 14.7 (8.3, 28.5) | 14.2 (10.4, 25.4) | 16.5 (11.9, 31.4) |
| $R_p$ | 1.7 (1.6, 1.9) | 1.8 (1.6, 2.1) | 1.6 (1.3, 2.2) | 1.7 (1.3, 2.5) |
| $\varphi$ | 0.79 (0.69, 0.89) | 0.90 (0.80, 0.96) | 0.91 (0.88, 0.95) | 0.92 (0.88, 0.97) |

Table S10: Parameters estimates with a contact matrix in MA. For each model tested (waning or leaky vaccine-derived immunity), the maximum likelihood estimates (95% CI) of trajectory matching (TM) and of the maximum iterated filtering algorithm (MIF) are presented. The best AIC value is indicated in boldface. Please refer to Table S5 for the meaning of the parameters.

### S3.4.3 Parameter estimates for a model integrating leakiness and waning

For definiteness, we examined the different modes of vaccine failure separately. In Table S11, we present the parameter estimates of a model mixing waning and leakiness. Notably, the likelihood of that model ($-3598.5$ (se: 1.8)) was almost identical to that of the waning model (–3596.4 (se: 1.5)), with confidence intervals for the leakiness enclosing 0. These results further confirm the lack of evidence for leakiness in the MA data.

| Quantity | Waning and leaky model | |
|---|---|---|
| | TM | MIF |
| $\log L$ | –3787.8 | –3598.5 (se: 1.8) |
| AIC | 7604 | **7225** |
| $\Delta$AIC | 379 | 0 |
| $q_1$ | 0.07 (0.05, 0.10) | 0.09 (0.05, 0.13) |
| $q_2$ | 0.03 (0.02, 0.06) | 0.04 (0.03, 0.11) |
| $q_3$ | 0.005 (0.004, 0.013) | 0.008 (0.004, 0.026) |
| $\rho_1(10+)$ | 0.88 (0.25, 1.00) | 0.74 (0.39, 1.00) |
| $\eta$ | 0.30 (0.18, 0.81) | 0.39 (0.04, 0.89) |
| $\rho_2(10+)$ | 0.26 (0.10, 0.53) | 0.29 (0.04, 0.67) |
| $\theta$ | 1.00 (0.40, 1.00) | 0.99 (0.35, 1.00) |
| $\tau$ | 0.56 (0.50, 0.61) | 0.22 (0.18, 0.28) |
| $\epsilon_A$ | 0.04 (0.02, 0.06) | 0.06 (0.03, 0.12) |
| $\alpha_V$ ($\mathrm{yr}^{-1}$) | 0.017 (0.006, 0.024) | 0.011 (0.002, 0.026) |
| $\epsilon_V$ | 0.00 (0.00, 0.04) | 0.01 (0.00, 0.03) |
| $R_0$ | 7.9 (5.8, 11.2) | 9.1 (5.3, 16.2) |
| $R_p$ | 1.7 (1.4, 1.8) | 1.8 (1.6, 2.0) |
| $\varphi$ | 0.80 (0.74, 0.89) | 0.83 (0.70, 0.92) |

Table S11: Parameter estimates of a model mixing waning and leakiness. For each model tested (waning or leaky vaccine-derived immunity), the maximum likelihood estimates (95% CI) of trajectory matching (TM) and of the maximum iterated filtering algorithm (MIF) are presented. The best AIC value is indicated in boldface. Please refer to Table S5 for the meaning of the parameters.

# References


1. C. D. Marchant, et al., *J Infect Dis* **169**(6), 1297 (1994).

2. T. H. Skoff, A. C. Cohn, T. A. Clark, N. E. Messonnier, S. W. Martin, *Arch Pediatr Adolesc Med* **166**(4), 344 (2012).

3. W. K. Yih, et al., *J Infect Dis* **182**(5), 1409 (2000).

4. 'Massachusetts Department of Public Health, School Immunization Data' (Accessed 5 August 2015).

5. C. G. Shapiro-Shapin, *Emerg Infect Dis* **16**(8), 1273 (2010).

6. J. S. Lavine, A. A. King, O. N. Bjørnstad, *Proc Natl Acad Sci U S A* **108**(17), 7259 (2011).

7. D. T. Karzon, *Postgrad Med J* **45**(520), 147 (1969).

8. J. Lavine, H. Broutin, E. T. Harvill, O. N. Bjørnstad, *Vaccine* **29**(1), 11 (2010).

9. M. Tanaka, et al., *JAMA* **290**(22), 2968 (2003).

10. J. D. Cryer, K.-s. Chan, *Time series analysis: with applications in R*. Springer, New York, 2nd ed edn. (2008).

11. US Census Bureau, 'Population estimates, accessed 16 December 2014'.

12. US Census Bureau, 'Population estimates, accessed 16 December 2014'.

13. M. Martinez-Bakker, K. M. Bakker, A. A. King, P. Rohani, *Proc Biol Sci* **281**(1783), 20132438 (2014).

14. H. J. Wearing, P. Rohani, *PLoS Pathog* **5**(10), e1000647 (2009).

15. J. C. Blackwood, D. A. T. Cummings, H. Broutin, S. Iamsirithaworn, P. Rohani, *Proc Natl Acad Sci U S A* **110**(23), 9595 (2013).

16. P. Rohani, X. Zhong, A. A. King, *Science* **330**(6006), 982 (2010).

17. R. Aguas, G. Gonçalves, M. G. M. Gomes, *Lancet Infect Dis* **6**(2), 112 (2006).

18. F. Magpantay, M. Riolo, M. Domenech de Cellès, A. King, P. Rohani, *SIAM J. Appl. Math.* **74(6)**, 1810 (2014).

19. F. M. G. Magpantay, M. Domenech DE Cellès, P. Rohani, A. A. King, *Parasitology* pp. 1–15 (2016).

20. M. A. Riolo, P. Rohani, *Proc Natl Acad Sci U S A* **112**(5), E472 (2015).

21. J. Mossong, et al., *PLoS Med* **5**(3), e74 (2008).

22. M. A. Riolo, A. A. King, P. Rohani, *Vaccine* **31**(49), 5903 (2013).

23. D. He, E. L. Ionides, A. A. King, *J R Soc Interface* **7**(43), 271 (2010).

24. J. D. Cherry, U. Heininger, in R. D. Feigin (ed.), *Feigin and Cherry's textbook of pediatric infectious diseases*, chap. 129, pp. 1616–1639. Saunders/Elsevier, Philadelphia, PA, 7th edn. (2014).

25. K. M. Edwards, M. D. Decker, in S. A. Plotkin, W. A. Orenstein, P. A. Offit (eds.), *Vaccines*, chap. 23, pp. 447–492. Elsevier Saunders, Philadelphia, Pa., 6th edn. (2013).

26. R. W. Sutter, S. L. Cochi, *JAMA* **267**(3), 386 (1992).

27. K. M. Farizo, et al., *Clin Infect Dis* **14**(3), 708 (1992).

28. S. G. Johnson, 'The nlopt nonlinear-optimization package' (?).

29. E. L. Ionides, D. Nguyen, Y. Atchadé, S. Stoev, A. A. King, *Proc Natl Acad Sci U S A* **112**(3), 719 (2015).

30. A. A. King, D. Nguyen, E. L. Ionides, *Journal of Statistical Software* **69**(1), 1 (2016).

31. C. T. Bauch, D. J. D. Earn, *Proc Biol Sci* **270**(1524), 1573 (2003).

32. S. L. Sheridan, R. S. Ware, K. Grimwood, S. B. Lambert, *JAMA* **308**(5), 454 (2012).

33. L. Fumanelli, M. Ajelli, P. Manfredi, A. Vespignani, S. Merler, *PLoS Comput Biol* **8**(9), e1002673 (2012).